

Genetic variability and population structure among wild Baird's tapirs

James E. Norton¹ and Mary V. Ashley

University of Illinois at Chicago, Department of Biological Sciences, (M/C 066), 845 West Taylor St., Chicago, IL 60607-7060, USA

¹ Current Address: Division of Pulmonary and Critical Care Medicine, Northwestern University Medical School, Tarry 14-707, 303 E. Chicago Ave., Chicago, IL 60611, USA

(Received 27 March 2003; accepted 1 November 2003)

Abstract

Baird's tapir, *Tapirus bairdii*, is an endangered, large, Central American mammal whose shy behaviour, solitary social structure and preference for deep tropical forests make it difficult to study using traditional field biology techniques. Despite great concern regarding its conservation status, very little is known about its population structure, mating system or dispersal patterns, information needed for designing appropriate management plans. Molecular genetic approaches can be of use in obtaining such information. In this study, we developed six polymorphic microsatellite genetic markers from genomic libraries of *T. bairdii* and *T. terrestris*. Using these markers, we completed the first assessment of genetic variability and population genetic structure in Baird's tapirs. Populations from southern Costa Rica and southern Panama were found to have low levels of genetic variability (allelic content, heterozygosity) but no indication of a recent population bottleneck. Bayesian and standard (F-statistic) analyses of genotype data indicate that the southern Costa Rican and southern Panamanian populations of Baird's tapirs were connected by some (at least one migrant per generation) gene flow prior to the destruction of intermittent habitat. These results indicate that the connection of these two populations by a MesoAmerican Biological Corridor (MBC) would be appropriate.

INTRODUCTION

Tapirs are an ancient genus of perissodactyl that was once both abundant and diverse across the globe. The family Tapiridae represents one of only three extant perissodactyl families (Schoch, 1989) and are the closest living relatives of the horses and the rhinoceroses. The only surviving genus is *Tapirus*, which is comprised of four extant species, *T. bairdii*, *T. terrestris*, *T. pinchaque* and *T. indicus*. Three of the species live in the Neotropics of Central and South America and the fourth exists in the tropical forests of Southeast Asia.

The Central American species, Baird's tapir, *T. bairdii*, once ranged continuously from southern Mexico to northwestern Ecuador, however it is now considered to be extinct in El Salvador and endangered in all other parts of its range (Fig. 1: IUCN, 1994). Habitat destruction and hunting are the primary causes for the decline in the number of tapirs (Brooks, Bodmer & Matola, 1997). With its entire geographical range being within the narrow peninsula of the Central American isthmus, the habitat of the Baird's tapir is at great risk of fragmentation as human development moves from the central highlands and Pacific

coast regions into the lowland tropics of the Atlantic coast (Jones, 1990; Brockett, 1998). Historically tapirs have been an important source of meat for the indigenous and rural people of Central America, however, recent published studies (Naranjo & Cruz, 1998; Flesher, 1999) suggest that hunting is on the decline. Nevertheless, hunting is still contributing to the decrease in numbers of tapirs.

The loss of tapirs to hunting and habitat destruction and the fragmentation and isolation of the remaining populations of Baird's tapirs increases their risk of extinction due to demographic, genetic and environmental stochasticity (Gilpin & Diamond, 1980; Caughley, 1994). Many surviving populations are isolated within protected habitats of parks and reserves. Unfortunately, only a few of these parks and reserves are believed to be large enough to sustain populations of tapirs for the long term (Flesher, 1999).

One suggested action to counter the problem of limited, small habitat patches is to connect fragmented populations with habitat corridors that facilitate migration and gene flow, and increase effective population size (Wilson & Willis, 1975; Noss, 1987; Beier & Noss, 1998). The establishment of such corridors may be particularly important for the long-term survival of this species.

However, tapirs have proven to be particularly difficult animals to study via traditional field biology techniques

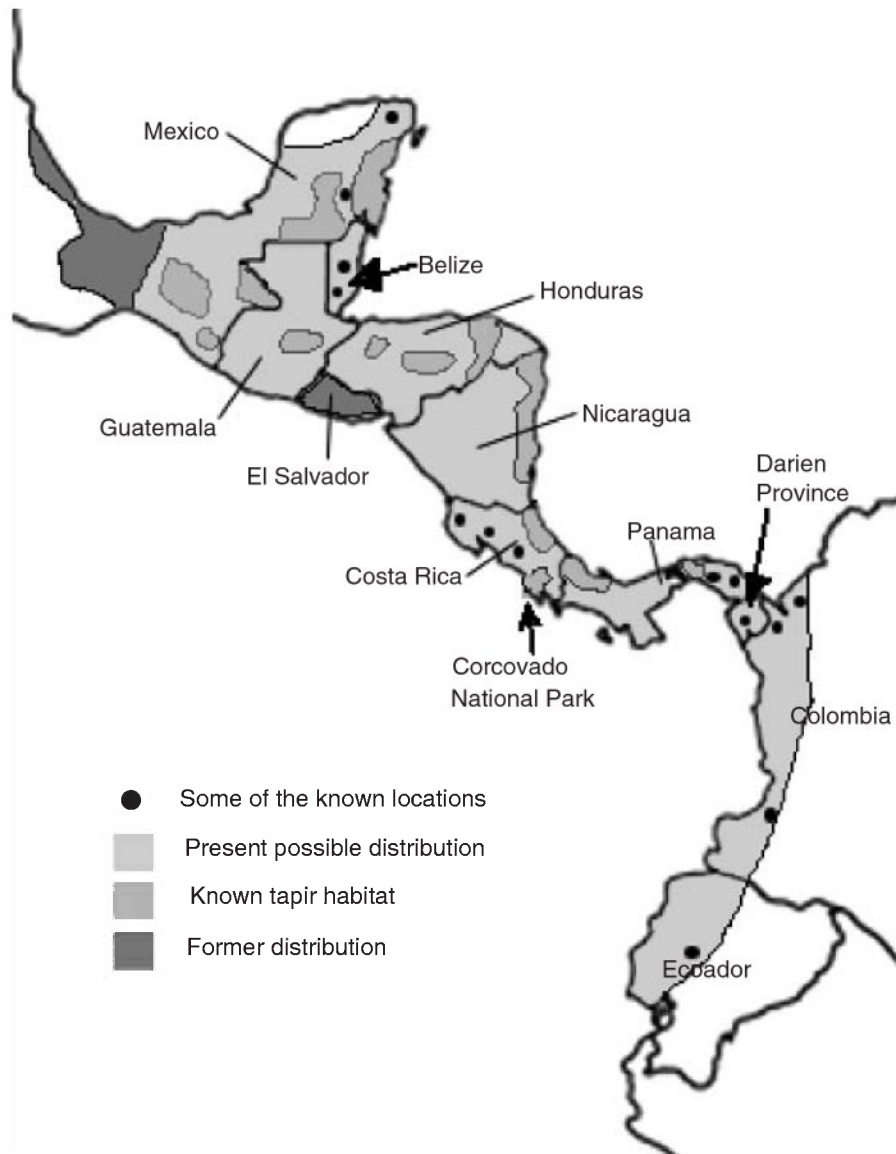


Fig. 1. Present and former distribution of *Tapirus bairdii*. Arrows indicate origin of wild population samples (modified from Brooks *et al.*, 1997).

and very little is known about their social structure, mating system or population structure, making it difficult to determine the priorities for establishing habitat connections. While radio-telemetry studies have been successful in exploring habitat usage, home ranges and daily behaviour patterns (Foerster & Vaughn, 2002), they have not been able to provide clear data on dispersal patterns. Population genetic data have often proven to be very useful for inferring levels of gene flow and population structure in animals that are difficult to observe or that travel long distances (Paetkau *et al.*, 1995; Andersen *et al.*, 1998; Feldheim, Gruber & Ashley, 2001; Sinclair *et al.*, 2001).

In this study, we developed six polymorphic microsatellite genetic markers from genomic libraries for *T. bairdii* and *T. terrestris*. A total of 33 DNA, tissue, hair or blood samples collected from wild populations of

Baird's tapirs from the Darien, Panama and Corcovado National Park, Costa Rica, as well as wild-caught animals in Central American zoos were genotyped at these loci. The microsatellite data were used to assess genetic variability and population structure of Baird's tapirs. Our objectives were to look for evidence of loss of genetic variation in remaining populations, and to infer historical patterns of connectivity and gene flow between populations that are currently isolated by habitat loss.

METHODS

Microsatellite development

The protocol described by Dow, Ashley & Howe (1995) was followed for the construction and screening of a Baird's tapir genomic library. Briefly, the genomic library

Table 1. Microsatellite loci developed from genomic libraries of *Tapirus terrestris* (*Tte*) and *T. bairdii* (*Tba*)

Locus	Motif	Primers	Annealing temperature (°C)	PCR product size	No. alleles
<i>Tte1</i>	(AC) ₂₅	Forward ATTAAGCAGATGCCAACCTGAAG Reverse CCCTGTGGTGTTTTTTGGATC	60	144–154	6
<i>Tte5</i>	(TC) ₁₀ (AC) ₁₀	Forward TCAACCCCTCCTGTCTCTTC Reverse ACTGATGGAACCGAGAAGCCAC	63	144–154	4
<i>Tte9</i>	(TC) ₁₁ (AC) ₁₀	Forward GCAGGCAAGGCTGAGGTTTT Reverse CTGCATTCTCCTTTGCCGACGGC	60	114–124	3
<i>Tte12</i>	(AC) ₁₉	Forward TTAGGGAAATAACAGGTCTGG Reverse GTTGTTTTGCATCCAAATTGG	55	162–170	3
<i>Tba15</i>	(AC) ₁₉	Forward TTGACCTTTTCATAAGCAGCC Reverse CCATCTCTCCATTCCAGTTC	53	214–238	7
<i>Tba23</i>	(AC) ₁₄	Forward ACAGTTTGTCCCTCCAAGTTG Reverse GCAGGTCAAATATACTGTCAGCCTGG	53	218–238	4
<i>Tba20</i>	(AC) ₁₅	Forward AACCCAAGTTGTCCGTCAACAG Reverse GCAGTTGTCTCTGACCGTGTGTTAG	65	246	1
<i>Tba21</i>	(AC) ₁₅	Forward ACCACTGCATATTTCTGCTGCC Reverse TCCTTACCTGCCTCCCAAGTAAAG	65	187	1
<i>Tba25</i>	(AC) ₁₇	Forward TCTTAGAAAATTAGTTAAGCCC Reverse CCATTATTCTGGTGATTACG	59	111	1

was screened for microsatellites using a cocktail of di-, tri- and tetra-nucleotide probes. Positive clones were sequenced (University of Illinois Core Facility) and polymerase chain reaction (PCR) primers were designed for inserts containing at least 10 dinucleotide repeats. Primers were synthesised and fluorescently labelled (Integrated DNA Technology, IDT). This same protocol was previously used to identify five microsatellite loci in a closely related species, *T. terrestris*, except that manual sequencing of insert DNA fragments using autoradiography was completed in the laboratory of M. V. Ashley. Microsatellite loci from both the *T. bairdii* and *T. terrestris* libraries were screened for polymorphism in *T. bairdii*.

Microsatellite genotypes were scored using an ABI 373a DNA Sequencer and analysed using ABI's GeneScan Fragment Analysis Software. PCR reactions were completed with 5–50 ng of DNA template added to 1.5–2.0 mM MgCl₂, 1.0 μM of each primer, 0.5 units of

Promega Taq, 2.0 μM bovine serum albumin, 1.0 μl of 10 × Promega PCR Buffer, 0.25 μM dNTP mix and sterile ddH₂O to a total volume of 10 μl (see Table 1 for annealing temperatures).

Sample collection

Tissue, blood and hair samples ($n = 33$) were collected from captive animals in zoos and conservation ranches as well as from a wild population in Corcovado National Park, Costa Rica ($n = 15$ were provided by C. Foerster and S. Hernandez-Divers, Dr. Vet. Mod.: Fig. 1). A sample of 15 wild animals from Panama was compiled from a combination of free ranging animals (5) and wild-caught captive animals (10). All samples are from animals from the Darien Region with the exception of one captive animal from near El Valle in central Panama. Finally, three samples were from wild-caught Belize animals held in captivity.

Analysis of genetic variability

An unweighted measure of alleles per locus (\bar{A}) was calculated for each Baird's tapir population as well as all population samples pooled together. We also used the computer programme FSTAT (Goudet, 1995) to calculate levels of allelic richness (R_s), which estimates alleles per locus independent of sample size. We then tested for significant differences between populations using a Wilcoxon signed ranks test. Because R_s adjusts measurements of alleles per locus to the smallest sample size (El Mousadik & Petit, 1996), we did not analyse the Belize sample ($n = 3$) independently but it was included in the pooled set of wild samples.

Observed and expected levels of heterozygosity (H_o , H_e) were calculated from the microsatellite genotype data using the computer programme Genetic Data Analysis (GDA: Lewis & Zaykin, 2001). Standard deviations for expected heterozygosity estimates within each population were also calculated using the square root of Weir's calculation of total variance (Weir, 1996: 146). Conformation of microsatellite genotype data to Hardy–Weinberg equilibrium (HWE) and linkage equilibrium was tested using the GenePop analysis package (Raymond & Rousset, 1995). Probabilities of HWE and genotypic disequilibrium were calculated using Fisher's exact test and GenePop's default analysis settings (Dememorization = 1000, Batches = 100, Iterations / batch = 1000). HWE was tested for each locus, each population and over all loci and populations for paired groups of populations, while genotypic disequilibrium was tested between all locus pairs within each population sample and within pooled samples of all populations. A sequential Bonferonni correction was applied to the data in order to compensate for the increased chance of a Type I error when conducting multiple significance tests (Rice, 1989).

We tested for the presence of null or non-amplifying alleles by estimating the frequencies of null alleles for loci having significant heterozygote deficiency following the methods described by Brookfield (1996). Using the expected frequency of the null allele an expected number of non-amplifying samples (homozygous for null allele) was calculated and a test for goodness-of-fit of the observed number of non-amplifying samples to the expected number of non-amplifying samples was completed using a simple χ^2 test.

Microsatellite genotype data were also tested for heterozygosity excess and shifts in allelic frequency distributions that would correlate with a recent genetic bottleneck using the programme Bottleneck (Piry, Luikart & Cornuet, 1999). We define the term 'recent' as being within the last several dozen generations, which in the case of Baird's tapirs would be approximately 75–150 years. This definition is based on the amount of time following a bottleneck event when an allele frequency shift would be detectable (Luikart *et al.*, 1998). Because less than 20 loci are being used in this analysis, the standard differences test was omitted and only the Wilcoxon signed ranks and sign test were used to obtain probability values for

excess levels of heterozygosity. All three proposed models of microsatellite mutation, Stepwise Mutation (SMM), Infinite Allele (IAM) and Two-Phase Models (TPM), were used to test for excess heterozygosity. TPM analysis was run with mutation model proportions of 95% SMM and 5% IAM as recommended by Piry *et al.* (1999).

Analysis of population structure

We assessed population structure using both standard (F-statistic) and Bayesian analyses of genotype data. In addition, allele frequency data were tested for differentiation between the Costa Rican and Panamanian populations. We used GDA (Lewis & Zaykin, 2001) to calculate Weir & Cockerham's (1984) version of Wright's (1978) F statistics and GenePop software was used to calculate Fisher exact test probabilities of allele frequency similarity across populations. The computer programme STRUCTURE (Pritchard, Stephens & Donnelly, 2000) was used to apply a Bayesian clustering approach for multi-locus genotyped data to test alternative models of population subdivision within the Panamanian and Costa Rican wild Baird's tapirs. The number of sub-populations (K) that is most probably present within the data set of microsatellite genotypes was determined by calculating the posterior probability of K populations for $K = 1, 2$ or 3 assuming a uniform prior probability. The natural log probabilities of sampling genotypes (X) from the pooled data of Costa Rican and Panamanian populations assuming K populations ($\text{Ln Pr}(X|K)$) was estimated by STRUCTURE using Markov Chain Monte Carlo methods. The log likelihoods were then used to calculate the posterior probability or the likelihood of each K using a form of Bayes' Theorem (Eqn 1: Pritchard *et al.*, 2000).

$$\text{Pr}(K) = e^{\text{Ln Pr}(X|K)} / [e^{\text{Ln Pr}(X|1)} + e^{\text{Ln Pr}(X|2)} + \dots + e^{\text{Ln Pr}(X|N)}] \quad (1)$$

where, N = largest number of sub-populations being tested.

Finally, STRUCTURE was used to evaluate the fidelity of each of the genotypes to their population sample (Costa Rica or Panama) by completing the maximum likelihood assignments with the known population data (Popinfo = ON) and the number of populations (Maxpops = $K = 2$) set to two.

RESULTS

Microsatellite development

Approximately 1200 recombinant bacterial colonies were screened. Of these, 43 colonies hybridised to simple sequence probes. Twenty-six of the 43 positive colonies were sequenced and seven were found to contain microsatellite repeats, resulting in an efficiency of 1:100 clones containing a microsatellite. Based on the number of clones found to contain microsatellites relative to the size of insert and number of colonies screened, microsatellites are estimated to occur within Baird's

Table 2. Average alleles/locus (\bar{A}), expected (H_e) and observed (H_o) levels of heterozygosity and inbreeding coefficient (f)

Population (n)	\bar{A}	H_e (+/- 2SD)	H_o	$f(F_{IS})^\dagger$
Costa Rica (15)	2.5	0.37 (+/- 0.08)	0.39	- 0.058
Panama (15)	3.33	0.43 (+/- 0.04)	0.41	0.032
All wild samples (33) [‡]	3.8	0.44 (+/- 0.19)	0.39	0.11

SD, standard deviation calculated from the square root of Weir's equation for total variance of heterozygosity (Eqn 4.6 in Weir, 1996).

[†] Fisher's exact test P values for heterozygote deficiency.

[‡] Includes three wild-caught samples from Belize Zoo.

tapir genomes approximately every 50–100 kilobases (((0.3–1kb/plasmid insert) \times 1200 plasmid inserts)/12 microsatellites), a frequency comparable to that reported in other mammalian species (Weber & May, 1989; Ellegren *et al.*, 1992; Ostrander *et al.*, 1992; Rooney *et al.*, 1999).

Primers were designed for five out of the seven microsatellite loci derived from the *T. bairdii* genome. Suitable primer sequences could not be designed for the other two loci due to insufficient sequence in the repeat flanking region. Additional sets of five primers were obtained from a *T. terrestris* genome. All 10 loci were screened for polymorphism by amplifying DNA samples from 10–15 Baird's tapir samples originating from different populations. Two of the five microsatellites identified in the *T. bairdii* genome were found to be polymorphic (4–7 alleles/locus) and were used for this study. Four out of the five *T. terrestris* loci were optimised and found to be polymorphic (3–6 alleles/locus) and were also used. PCR amplification of all six microsatellite loci was attempted on all DNA samples. Four (2%) out of 198 (33 samples \times 6 loci) PCR reactions failed to produce sufficient product for genotyping.

Genetic variability

Allelic diversity (\bar{A}) ranged from 2.5–3.33 alleles/locus ($n = 15$), while \bar{A} in the pooled sample of all wild animals was 3.8 alleles/locus ($n = 33$) (Table 2). The allelic richness (R_s) across loci in Baird's tapir populations ranged from 1.0–5.30 alleles/locus (Table 3) but did not differ significantly between population samples ($P > 0.05$). Each wild population sample, with the exception of the Belize sample ($n = 3$), contained private alleles (Table 4), but these were present at low frequencies and it is possible that these alleles are present in the other populations but were not sampled.

Observed and expected heterozygosity were found to be in the lower range of microsatellite heterozygosity among mammalian species (0.39–0.41 and 0.37–0.44, respectively) and were not significantly different between Costa Rica and Panama (Table 2). Tests for deviations from Hardy–Weinberg equilibrium showed that none of the loci demonstrated significant levels of heterozygote excess at any locus before Bonferonni correction ($P > 0.05$). In the pooled sample of all wild animals, locus *Tte1*

Table 3. Allelic richness (R_s) per locus and population based on minimum sample size of 10 diploid individuals

Locus	Population		
	Costa Rica	Panama	All wild
<i>Tte1</i>	1.667	1.998	2.150
<i>Tte5</i>	1.992	3.333	2.605
<i>Tte9</i>	2.000	1.000	1.978
<i>Tte12</i>	2.998	2.897	2.998
<i>Tba15</i>	3.666	5.300	4.29
<i>Tba23</i>	2.000	3.837	3.33

Table 4. Allelic content and frequency in wild and captive populations of *Tapirus bairdii*

Locus	$P(\alpha)^{6,5=0.008,0.01}$	Alleles	Belize ($n = 3$)	Costa Rica ($n = 15$)	Panama ($n = 15$)
<i>Tte1</i>	0.051	144	–	–	0.17
		146	1.0	0.97	0.83
		154	–	0.03	–
<i>Tte5</i>	0.006	144	1.0	0.87	0.50
		148	–	–	0.035
		150	–	–	0.035
<i>Tte9</i>	0.026	154	–	0.13	0.43
		114	0.5	0.20	–
		124	0.5	0.80	1.0
<i>Tte12</i>	0.032	162	–	0.28	0.43
		168	1.0	0.44	0.07
		170	–	0.28	0.50
<i>Tba15</i>	0.010	214	–	–	0.035
		218	–	0.53	0.23
		222	0.67	0.235	0.12
		224	–	–	0.035
		226	–	–	0.035
		228	0.33	0.20	0.54
		238	–	0.035	–
<i>Tba23</i>	0.036	218	0.83	0.20	0.07
		234	–	–	0.10
		236	0.17	0.80	0.73
		238	–	–	0.10

[†] Fisher exact test probabilities for similarity in allele frequencies between Cost Rican and Panamanian populations.

Bold type indicates private alleles within the wild population or significant P value.

showed a significant heterozygote deficit before, but not after, sequential Bonferonni correction. Locus *Tte12* and locus *Tba23* did demonstrate significant deviations from expected genotypic frequencies ($P = 0.000$) suggesting that the two loci are linked or segregating together. However, independent calculations of F statistics excluding these loci did not greatly alter the results and the qualitative conclusions remained the same. Consequently, all loci data were retained and used in estimating population structure.

Heterozygote deficiency can be caused by the presence of a null or non-amplifying allele at a locus. The expected number of non-amplifying individuals in each sample set

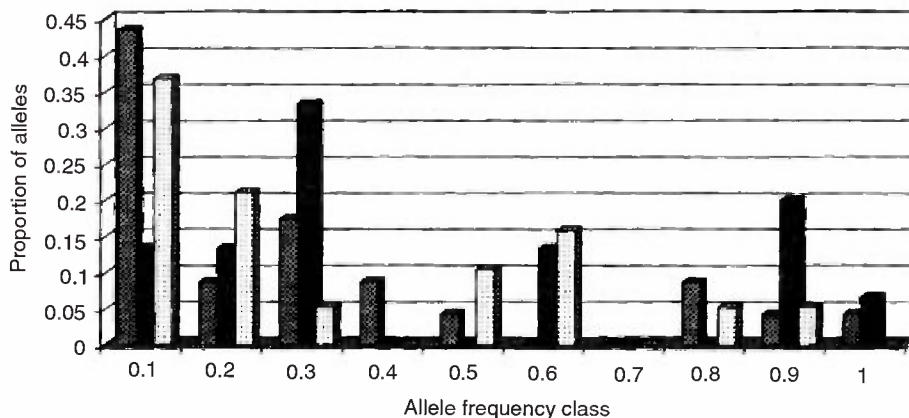


Fig. 2. Allele frequency distributions for all loci genotyped. The values on the x-axis represent the frequencies of alleles at each loci. The values along the y-axis indicate the proportion of all alleles within each frequency class. ▨, all wild populations; ■, Corcovado National Park, Costa Rica; □, Darien, Panama.

calculated from the estimation of the frequency of null alleles present in the sample was not significantly different from the observed number of non-amplifying samples ($P > 0.05$).

The results of tests for heterozygosity excess under all models, IAM, TPM and SMM, provided no evidence for recent population bottlenecks in either population. Neither population deviated significantly from mutation-drift equilibrium levels of heterozygosity (IAM: Sign Test, $P = 0.14$, Wilcoxon two tailed, $P = 0.15$; TPM: Sign Test, $P = 0.21$, Wilcoxon two tailed, $P = 0.21$; SMM: Sign Test, $P = 0.05$, Wilcoxon two tailed, $P = 0.07$). However, the allele frequency distribution in the Costa Rican sample displayed a mode shift to the right, indicating a putative bottleneck may have occurred in this population (Fig. 2). The other population samples produced a normal L-shaped distribution (Fig. 2).

F-statistics and genic differentiation

Based on Wright's interpretation of F_{ST} values (Wright, 1978), theta (θ , F_{ST}) values (0.114) and 95% bootstrap confidence intervals (0.059–0.18) between the Panamanian and Costa Rican wild populations can be interpreted as indicating the presence of moderate levels of population differentiation. The F value (0.075), or the inbreeding coefficient that reflects the impact of both non-random mating within subpopulations and genetic drift between subpopulations, and the corresponding 95% confidence intervals ($F = -0.213$ – 0.245), is not significantly different from zero indicating that the observed level of differentiation would be primarily a result of drift within each population as opposed to non-random mating. Additionally, the low inbreeding coefficient (f) within the two populations (CR, PA: Table 2) provides further evidence that the observed differentiation between population samples is primarily a result of genetic drift and not due to increased inbreeding. Tests for differences in allele frequencies between the Costa Rican

and Panamanian population samples found that allele frequencies at two of the six loci were significantly different after Bonferonni correction (Table 4).

Bayesian analysis

The posterior probabilities calculated from the Costa Rican and Panamanian data indicate that the samples from Costa Rica and Panama are more likely to have come from a single panmictic population ($P = 0.999$) versus two ($P \approx 0$) or three ($P \approx 0$) sub-populations. In addition, assignment test results assigned almost all individuals to both the Costa Rican and the Panamanian population with nearly equal likelihood. One recently published study (Maudet *et al.*, 2002) indicates that the methods used in the programme STRUCTURE have reduced effectiveness in correctly assigning individuals to populations unless population differentiation (F_{ST}) is considerable.

DISCUSSION

Microsatellites were found to be present within the Baird's tapir genome at frequencies comparable to those observed in other mammalian species such as humans (Weber & May, 1989), horses (Ellegren *et al.*, 1992), dogs (Ostrander *et al.*, 1992), and bowhead whales (Rooney *et al.*, 1999). However, the microsatellites identified in Baird's tapirs proved to have low levels of polymorphism relative to what is expected within mammalian species. A survey of microsatellite loci in the EMBL and GenBank databases revealed that microsatellites with 12 or more dinucleotide repeats were almost always polymorphic even when amplified across species (Moore *et al.*, 1992). Yet among the five microsatellite loci larger than 12 dinucleotide repeats isolated from the Baird's tapir genome, three were found to be monomorphic when screened across 10 unrelated individuals from across the range of Baird's tapirs. Because of the high proportion of monomorphic loci present in the *T. bairdii*

Table 5. Levels of microsatellite variability and population structure in large mammals

	<i>N</i>	<i>n</i>	<i>H_e</i>	<i>H_o</i>	<i>F_{ST}</i>	<i>L</i>	<i>A</i>
Asian water buffalo (<i>Bubalus bubalis</i>): Barker <i>et al.</i> (1997)	200	25	0.48	0.51	0.168	21	7.1
Atlantic walrus (<i>Odobenus rosmarus</i>): Andersen <i>et al.</i> (1998)	105	26.5	0.61	0.65	0.058	11	5.6–6.0
Barbary Macaques (<i>Macaca sylvanus</i>): Von Segesser <i>et al.</i> (1999)	159	17.6	0.65	0.32–0.77	0.118	6	5.16
Black-Footed rock-wallaby (<i>Petrogale lateralis</i>): Eldridge <i>et al.</i> (1999)							
(mainland)	22.0	22.0	0.562	–	–	10	4.4
(island bottleneck)	14.0	14.0	0.053	–	–	10	1.2
Cougar (<i>Puma concolor</i>): Sinclair <i>et al.</i> (2001)	50.0	5.0	0.655	0.626	0.039	9	4.44
Koala (<i>Phascolarctos cinereus</i>): Houlden <i>et al.</i> (1996)							
(Non-bottleneck)	80	–	0.851	–	0.09–0.16	6	11.5
(Bottleneck)	160	–	0.436	–	0.03–0.21	–	5.3
Moose (<i>Alces alces</i>): Broders <i>et al.</i> (1999)							
(Source)	208.0	34.6	–	0.41	–	5	3.4
(Founded)	263.0	65.8	–	0.22–0.31	–	5	2.6
Baird's tapir (<i>Tapirus bairdii</i>)	33.0	15.0	0.37–0.43	0.39–0.41	0.059–0.18	6	2.5–3.33
Spanish Celtic horse (<i>Caballus caballus</i>): Canon <i>et al.</i> (2000)	541	20.0	0.78	0.72	0.078	13	5.8
Swayne's hartebeest: (<i>Alcelaphus buselaphus swaynei</i>): Flagstad <i>et al.</i> (2000)	44.0	20.0	0.70	0.64	–	8	5.0
Grey wolf (<i>Canis lupus</i>): Forbes & Boyd (1997)	172.0	39.7	0.64	0.59	0.168	9	4.1

N, total sample size; *n*, average population sub-sample size; *H_e*; expected heterozygosity; *H_o*, observed heterozygosity; *L*, number of loci; *A*, average alleles/locus.

genome, enrichment protocols (Ostrander *et al.*, 1992; Karagyozov, Kalcheva & Chapman, 1993; Hamilton *et al.*, 1999; Paetkau, 1999) are recommended to increase the concentration of microsatellite fragments present in the ligation solution.

The samples of Baird's tapir populations assessed in this study were found to have levels of heterozygosity and allelic diversity that are among the lowest reported in large mammals (Table 5). Moreover, the levels of heterozygosity in Baird's tapirs are also more consistent with the levels in populations of large mammals with a history of isolation or bottlenecks (Table 5). Low levels of genetic variability, heterozygosity and allelic diversity have been used as evidence of genetic bottlenecks in populations that are known to have experienced severe population declines (e.g. Houlden *et al.*, 1996; LePage *et al.*, 2000; Goodman *et al.*, 2001; Whitehouse & Harley, 2001; Williams *et al.*, 2002).

The Baird's tapirs sampled here do not demonstrate the heterozygosity excess expected in populations that have experienced a recent bottleneck, refuting the occurrence of a recent bottleneck or founder event in the Corcovado, Costa Rica or Panamanian populations. However, the allele frequency mode shift observed in the Corcovado population conflicts with the lack of heterozygote excess and is consistent with a population bottleneck (Fig. 2).

The inability of tests for heterozygosity excess and allele frequency mode shifts to uniformly detect the impact of a recent bottleneck suggests that any putative bottleneck would have had to have occurred at least 50–100 years (25–50 generations) ago. This time frame is consistent with the beginning of major industrial and agricultural development in Central American countries (Jones, 1990; Brockett, 1998) and this development would have resulted in the destruction and fragmentation of tapir habitat as well as increased hunting of tapirs by farmers, ranchers and developers. The occurrence of a bottleneck event in the more distant past would significantly decrease the number of alleles in the species. Then, several generations of increased inbreeding created by the rapid decline in effective population size would reduce heterozygosity until reaching a drift/mutation equilibrium state consistent with the number of alleles that remain in the population. However, the small sample size and number of loci used in this analysis limits the power of the BOTTLENECK tests applied and further sampling and analysis with an expanded set of loci is necessary before definitive conclusions can be drawn.

These results are consistent with the population genetic modelling described by Nei, Maruyama & Chakraborty (1975), who predicted that populations that are not able

to recover population size quickly following a bottleneck will experience a greater loss of genetic variation and take longer to recover heterozygosity. Furthermore, the mutation process to introduce new alleles would take thousands of generations even for rapidly mutating microsatellites (10^{-3} mutations/locus/generation = 10^3 generations/mutation/locus) resulting in the maintenance of low allelic diversity for thousands of years following a bottleneck. In addition, the small numbers of alleles per locus after a severe bottleneck would cause mutation-drift equilibrium heterozygosity to stabilise at a low level and increase slowly as new alleles are introduced through mutation and increase in frequency through drift (Nei *et al.*, 1975).

Recent population genetic studies (e.g. LePage *et al.*, 2000; Waits *et al.*, 2000; Goodman *et al.*, 2001; Whitehouse & Harley, 2001; Williams *et al.*, 2002) of populations that are known to have experienced population bottlenecks support these theoretical conclusions. For example Scandinavian brown bears known to have experienced a severe population bottleneck during the nineteenth century as a result of vigorous extermination program did not reveal a significant loss of heterozygosity or allelic diversity when compared to non-bottlenecked populations of brown bears (Waits *et al.*, 2000). Waits *et al.* (2000) suggested one explanation for these results is that the population bottleneck occurred in four discrete geographical regions in Scandinavia, possibly allowing four independent genetic drift effects to randomly preserve different combinations of alleles within each sub-population. Then, because most of the bear's habitat remained relatively intact, rapid population growth and reestablishment of gene flow with the other sub-populations preserved both allelic diversity and heterozygosity. Alternatively, populations of elk (Williams *et al.*, 2002), wallaby (LePage *et al.*, 2000) and African elephants (Whitehouse & Harley, 2001) that are known to have experienced severe population bottlenecks, but also experienced restricted population growth and gene flow due to loss of habitat and continued pressure from hunting, showed the expected extremely low levels of genetic variability.

One microsatellite locus exhibited heterozygote deficiency. Although it has been shown to be extremely difficult to isolate mechanisms acting to reduce heterozygosity (Christiansen *et al.*, 1974), the evidence in this case supports the presence of a null allele as the cause of heterozygote deficiency in locus *Tte1*. One DNA sample that was easily scorable at the five other loci failed to produce bands at locus *Tte1*, consistent with the presence of a null allele. Also, the expected number of non-amplifying individuals in each sample set calculated from the estimation of the frequency of null alleles present in the sample was not significantly different from the observed number of non-amplifying samples.

The fragmentation of the tropical forest habitat of the tapir by its conversion to grasslands for pastoral and agricultural development could result in the fragmentation and isolation of formerly connected tapir populations. Behavioural and ecological studies of Baird's tapirs have

revealed their avoidance of open grasslands such as those created for agriculture and resulting in potential barriers to tapir movement (Flesher & Ley, 1996; Naranjo & Cruz, 1998; Flesher, 1999). Furthermore, tapirs appear to alter their behaviour to avoid humans and the settlement of tapir habitat may not only deter their movement through human populated areas but may also create a sink due to the increased risk of mortality due to hunting (Barongi, 1993; Flesher & Lay, 1996; Flesher, 1999; Foerster & Vaughan, 2002). Although the Baird's tapir populations sampled here have experienced recent population disturbances (habitat destruction, population decline, founder events for captive populations) and, therefore, are unlikely to provide an accurate estimate of historic patterns of gene flow, population structure analysis should still provide a measurement of broad population structure patterns and levels of genetic differentiation. Furthermore, qualitative assessments of the occurrence or absence of past gene flow can be inferred.

The microsatellite data analysed in this study suggest the presence of at least moderate levels of population structure between the populations of Baird's tapirs in Corcovado National Park, Costa Rica and animals from the Darien, Panama. Populations that have similar allele frequencies at loci belong in the same Evolutionarily Significant Units (ESUs) while those that are highly divergent have some barrier to gene flow and migration and are most probably not in the same Management Units (MU) and may not belong within the same ESU (Moritz, 1994, 1999; Bowen, 1998; Forbes & Hogg, 1999). The similarities in microsatellite loci content, moderate levels of differentiation at frequency levels (Table 4) and Bayesian analytical pooling of the two population samples into a single population supports placing these two populations within the same ESU. However, the moderate levels of differentiation ($\theta = 0.114$) indicate a historic pattern of reduced gene flow between the two populations. The indicated level of genetic drift as well as current population isolation due to habitat destruction support treating the two populations as independent MUs.

No clear means of estimating the likelihood or severity of outbreeding depression that may result from facilitating the mixing of two populations from genetic differentiation parameters such as F_{ST} have been established (Hedrick & Miller, 1992; Frankham, 1995). However, it is more advisable to link or exchange individuals between populations whose level of structure indicates at least minimal levels of historic gene flow ($F_{ST} < 0.2$: Forbes & Hogg, 1999). Mills & Allendorf (1996) demonstrated, through modelling, that F_{ST} values for populations that are exchanging one migrant per generation will reach an equilibrium value of 0.2. It can thus be inferred that populations where $F_{ST} \leq 0.2$ have historical migration rates of at least one migrant per generation. Based on our assessment of population structure, Baird's tapirs in Corcovado National Park, Costa Rica and the Darien Region of Panama were historically linked by minimal levels of gene flow (one migrant per generation). Restoring connectivity between these two populations by connecting these regions via a Mesoamerican Biological Corridor

(MBC) is appropriate and would re-establish what appear to be historic connections.

Acknowledgements

We thank the many individuals and institutions who were extremely helpful in the collection of samples for this study: Charles Foerster, Sonia Hernandez-Divers, Manuel Ruiz-Garcia, Belize Zoo, El Nispero Zoo, Summit Zoo, Villa de Griselda Conservation Ranch and Jaime Padilla Vellis and El Siglo Newspaper. We also thank two anonymous referees for their helpful comments. This work was funded in part by the American Society of Mammalogists, Chicago Zoological Society and Sophie Danforth Conservation Fund. Baird's tapir samples were imported from Belize and Panama under CITES permits 99US017651/9 and 01US028420/9, respectively. This research was completed in partial fulfillment of the requirements for the Masters degree (J.E.N.) at the Graduate College of the University of Illinois at Chicago.

REFERENCES

- Andersen, L. W., Born, E. W., Gjertz, I., Wiig, O., Holm, L.-E. & Bendixen, C. (1998). Population structure and gene flow of the Atlantic walrus (*Odobenus rosmarus rosmarus*) in the Eastern Atlantic Arctic based on mitochondrial DNA and microsatellite variation. *Mol. Ecol.* **7**:1323–1336.
- Barker, J. S. F., Moore, S. S., Hetzel, D. J. S., Evans, D., Tan, S. G. & Byrne, K. (1997). Genetic diversity of Asian water buffalo (*Bubalus bubalis*): microsatellite variation and a comparison with protein-coding loci. *Anim. Genet.* **28**: 103–115.
- Barongi, R. A. (1993). Husbandry and conservation of tapirs. *Int. Zoo Yb.* **32**: 7–15.
- Beier, P. & Noss, R. F. (1998). Do habitat corridors really provide connectivity? *Conserv. Biol.* **12**: 1241–1252.
- Bowen, B. (1998). What is wrong with ESUs?: The gap between evolutionary theory and conservation principles. *J. Shellfish Res.* **17**: 1355–1358.
- Brockett, C. D. (1998). *Land power and poverty: agrarian transformation and political conflict in Central America*. Boulder, CA: Westview Press.
- Broders, H. G., Mahoney, S. P., Montevocchi, W. A. & Davidson, W. S. (1999). Population genetic structure and the effect of founder events on the genetic variability of moose, *Alces alces*, in Canada. *Mol. Ecol.* **8**: 1309–1315.
- Brooks, D. M., Bodmer, D. M. & Matola, S. (Eds). (1997). *Status survey and conservation action plan-Tapirs*. IUCN/SSC Tapir Specialist Group-Action Plan. Gland, Switzerland: IUCN.
- Brookfield, J. F. Y. (1996). A simple new method for estimating null allele frequency from heterozygote deficiency. *Mol. Ecol.* **5**: 453–455.
- Cañon, J., Checa, M. L., Carleos, C., Vega-Pla, J. L., Vallejo, M. & Dunner, S. (2000). The genetic structure of Spanish Celtic horse breeds inferred from microsatellite data. *Anim. Genet.* **31**: 39–48.
- Caughley, G. (1994). Directions in conservation biology. *J. Anim. Ecol.* **63**: 215–244.
- Christiansen, F. B., Frydenberg, O., Gyldenholm, A. O. & Simonsen, V. (1974). Genetics of *Zoarces* populations VI. Further evidence, based on age group samples, of a heterozygote deficit in the ESTIII polymorphism. *Hereditas* **77**: 225–236.
- Dow, B. D., Ashley, M. V. & Howe, H. F. (1995). Characterization of highly variable (GA/CT)_n microsatellites in the bur oak, *Quercus macrocarpa*. *Theoret. Appl. Genet.* **91**: 137–141.
- Eldridge, M. D. B., King, J. M., Loupis, A. K., Spencer, P. B. S., Taylor, A. C., Pope, L. C. & Hall, G. P. (1999). Unprecedented low levels of genetic variation and inbreeding depression in an island population of the black-footed rock-wallaby *Petrogale lateralis*. *Conserv. Biol.* **13**: 531–541.
- Ellegren, H., Johansson, M., Sandberg, K. & Andersson, L. (1992). Cloning of highly polymorphic microsatellites in the horse. *Anim. Genet.* **23**: 133–142.
- El Mousadik, A. & Petit, R. J. (1996). High level of genetic differentiation for allelic richness among populations of the argan tree (*Argania spinosa* (L.) Skeels) endemic to Morocco. *Theoret. Appl. Genet.* **92**: 832–839.
- Feldheim, K. A., Gruber, S. H. & Ashley, M. V. (2001). Population genetic structure of the lemon shark (*Negaprion brevirostris*) in the western Atlantic: DNA microsatellite variation. *Mol. Ecol.* **10**: 295–303.
- Flagstad, O., Syvertsen, P. O., Stenseth, N. C., Stacy, J. E., Olsaker, I., Roed, K. H. & Jakobsen, K. S. (2000). Genetic variability in Swayne's hartebeest, an endangered antelope of Ethiopia. *Conserv. Biol.* **14**: 254–264.
- Flesher, K. (1999). Preliminary notes on the conservation status of Baird's tapir *Tapirus bairdii* in north-eastern Honduras. *Oryx* **33**: 294–300.
- Flesher, K. & Ley, E. (1996). A frontier model for landscape ecology: the tapir in Honduras. *Environ. Ecol. Stat.* **3**: 119–125.
- Foerster, C. R. & Vaughn, C. (2002). Home range, habitat use, and activity of Baird's tapir in Costa Rica. *Biotropica* **34**: 423–437.
- Forbes, S. H. & Hogg, J. T. (1999). Assessing population structure at high levels of differentiation microsatellite comparisons of bighorn sheep and large carnivores. *Anim. Conserv.* **2**: 223–233.
- Forbes, S. H. & Boyd, D. K. (1997). Genetic structure and migration of native and re-introduced Rocky Mountain wolf populations. *Conserv. Biol.* **11**: 1226–1234.
- Frankham, R. (1995). Conservation genetics. *Annu. Rev. Genet.* **29**: 305–327.
- Gilpin, M. E. & Diamond, J. M. (1980). Subdivision of nature reserves and the maintenance of species diversity. *Nature* **285**: 567–568.
- Goudet, J. (1995). FSTAT (Version 1.2): a computer program to calculate F-statistics. *J. Hered.* **86**: 485–486.
- Goodman, S. J., Tamate, H. B., Wilson, R., Nagata, J., Tatsuzawa, S., Swanson, G. M., Pemberton, J. M. & McCullough, D. R. (2001). Bottlenecks, drift and differentiation: the population structure and demographic history of sika deer (*Cervus nippon*) in the Japanese archipelago. *Mol. Ecol.* **10**: 1357–1370.
- Hamilton, M. B., Pincus, E. L., Di-Fiore, A. & Fleischer, R. C. (1999). Universal linker and ligation procedures for construction of genomic DNA libraries enriched for microsatellites. *Biotechniques* **27**: 500–507.
- Hedrick, P. W. & Miller, P. S. (1992). Conservation genetics: techniques and fundamentals. *Ecol. Applic.* **2**: 30–46.
- Houlden, B. A., England, P. R., Taylor, A. C., Greville, W. D. & Sherwin, W. B. (1996). Low genetic variability of the koala *Phascolarctos cinereus* in south-eastern Australia following a severe population bottleneck. *Mol. Ecol.* **5**: 269–281.
- IUCN (1994). *The IUCN mammal red data book, part 1: IUCN red list of threatened animals*. Gland, Switzerland: World Union for Nature.
- Jones, J. R. (1990). *Colonization and environment: land settlement projects in Central America*. Tokyo: United Nations University Press.
- Karagoyzov, L., Kalcheva, I. D. & Chapman, V. M. (1993). Construction of random small-insert genomic libraries highly enriched for simple sequence repeats. *Nucl. Acid Res.* **21**: 3911–3912.
- LePage, S. L., Livermore, R. A., Cooper, D. W. & Taylor, A. C. (2000). Genetic analysis of a documented population bottleneck: introduced Bennet's wallabies (*Macropus rufogriseus rufogriseus*) in New Zealand. *Mol. Ecol.* **9**: 753–763.
- Lewis, P. O. & Zaykin, D. (2001). *Genetic Data Analysis*. Computer program for the analysis of allelic data. Version 1.0 (d16c). Free program distributed by the authors over the internet from <http://lewis.eeb.uconn.edu/lewishome/software.html>.

- Luikart, G., Allendorf, F. W., Cornuet, J.-M. & Sherwin, W. B. (1998). Distortion of allele frequency distributions provides a test for recent population bottlenecks. *J. Hered.* **89**: 238–247.
- Maudet, C., Miller, C., Bassano, B., Breitenmoser-Wursten, C., Gauthier, D., Obexer-Ruff, G., Michallet, J., Taberlet, P. & Luikart, G. (2002). Microsatellite DNA and recent statistical methods in wildlife conservation management: applications in Alpine ibex (*Capra ibex*). *Mol. Ecol.* **11**: 421–436.
- Mills, L. S. & Allendorf, F. W. (1996). The one-migrant-per-generation rule in conservation and management. *Conserv. Biol.* **10**: 1509–1518.
- Moore, S. S., Barendse, W., Berger, K. T., Armitage, S. M. & Hetzel, D. J. S. (1992). Bovine and ovine DNA microsatellites from EMBL and GenBank Databases. *Anim. Genet.* **23**: 463–467.
- Moritz, C. (1994). Defining 'evolutionary significant units' for conservation. *Trends Ecol. Evol.* **9**: 373–375.
- Moritz, C. (1999). Conservation units and translocations: strategies for conserving evolutionary processes. *Hereditas.* **130**: 217–228.
- Naranjo, P. E. J. & Cruz, A. (1998). Ecology of tapirs (*Tapirus bairdii*) in the Sepultura Biosphere Reserve Chiapas, Mexico. *Acta Zoo. Mex.* **73**: 111–125.
- Nei, M., Maruyama, T. & Charkraborty, R. (1975). The bottleneck effect and genetic variability in populations. *Evolution* **29**: 1–10.
- Noss, R. F. (1987). Corridors in real landscape: a reply to Simberloff and Cox. *Conserv. Biol.* **1**: 159–164.
- Ostrander, E. A., Jong, P. M., Rine, J. & Duyk, G. (1992). Construction of small-insert genomic DNA libraries highly enriched for microsatellite repeat sequences. *Proc. Natl. Acad. Sci. USA* **89**: 3419–3423.
- Paetkau, D. (1999). Microsatellites obtained using strand extension: an enrichment protocol. *Biotechniques* **26**: 690–697.
- Paetkau, D., Calvert, W., Stirling, I. & Strobeck, C. (1995). Microsatellite analysis of population structure in Canadian polar bears. *Mol. Ecol.* **4**: 347–354.
- Piry, S., Luikart, G. & Cornuet, J.-M. (1999). BOTTLENECK: a computer program for detecting recent reductions in the effective population size using allele frequency data. *J. Hered.* **90**: 502–503.
- Pritchard, J. K., Stephens, M. & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics* **155**: 954–959.
- Raymond, M. & Rousset, F. (1995). Genepop (version 1.2): population genetics software for exact tests and ecumenicism. *J. Hered.* **86**: 248–249. (Software available at: <http://wbiomed.curtin.edu.au/genepop>.)
- Rice, W. R. (1989). Analyzing tables of statistical tests. *Evolution* **43**: 223–225.
- Rooney, A. P., Honeycutt, R. L., Davis, S. K. & Derr, J. M. (1999). Evaluating a putative bottleneck in a population of bowhead whales from patterns of microsatellite diversity and genetic disequilibria. *J. Mol. Ecol.* **49**: 682–690.
- Schoch, R. M. (1989). A review of Tapiroids. In: *The evolution of Perissodactyls*: 298–320. Prothero, D. R. & R. M. Schoch (Eds). New York: Oxford University Press.
- Sinclair, E. A., Swenson, E. L., Wolfe, M. L., Choate, D. C., Bates, B. & Crandall, K. A. (2001). Gene flow estimates in Utah's cougars imply management beyond Utah. *Anim. Conserv.* **4**: 257–264.
- Von Segesser, F., Menard, N., Gaci, B. & Martin, R. D. (1999). Genetic differentiation within and between Algerian subpopulations of Barbary macaques (*Macaca sylvanus*): evidence from microsatellites. *Mol. Ecol.* **8**: 433–442.
- Waits, L., Taberlet, P., Swenson, J. E., Sandgren, F. & Franzen, R. (2000). Nuclear DNA microsatellite analysis of genetic diversity and gene flow in the Scandinavian brown bear (*Ursus arctos*). *Mol. Ecol.* **9**: 421–431.
- Weber, J. L. & May, P. E. (1989). Abundant class of human DNA polymorphisms, which can be typed using the polymerase chain reaction. *Am. J. Hum. Genet.* **44**: 388–396.
- Weir, B. S. (1996). *Genetic Data Analysis II: methods for discrete population genetic data*. Sunderland, MA: Sinauer Assoc. Inc.
- Weir, B. S. & Cockerham, C. C. (1984). Estimating F-statistics for the analysis of population structure. *Evolution* **38**: 1358–1370.
- Whitehouse, A. M. & Harley, E. H. (2001). Post-bottleneck genetic diversity of elephant populations in South Africa, revealed using microsatellite analysis. *Mol. Ecol.* **10**: 2139–2149.
- Williams, C. L., Serfass, T. L., Cogan, R. & Rhodes, O. E. (2002). Microsatellite variation in the reintroduced Pennsylvania elk herd. *Mol. Ecol.* **11**: 1299–1310.
- Wilson, E. O. & Willis, E. O. (1975). Applied biogeography. In: *Ecology and Evolution of Communities*: 522–534. Cody, M. L. & Diamond, J. M. (Eds). Cambridge, MA: Harvard University Press.
- Wright, S. (1978). *Evolution and the genetics of population. Vol. 4: variability within and among natural populations*. Chicago: University of Chicago Press.